

Shin Lin CHEN

1633 CHEN

H8

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/284,180DATE: 08/17/1999
TIME: 17:17:14

Input Set: I284180.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

P.S.

ENTERED

```
1 <110> APPLICANT: KIMURA, Toru
2 KIKUCHI, Kaoru
3 <120> TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN W
4 <130> FILE REFERENCE: 20-4546P
5 <140> CURRENT APPLICATION NUMBER: US/09/284,180
6 <141> CURRENT FILING DATE: 1999-06-09
7 <160> NUMBER OF SEQ ID NOS: 14
8 <170> SOFTWARE: PatentIn Ver. 2.0
9 <210> SEQ ID NO 1
10 <211> LENGTH: 4008
11 <212> TYPE: DNA
12 <213> ORGANISM: Rattus norvegicus
13 <220> FEATURE:
14 <221> NAME/KEY: 5'UTR
15 <222> LOCATION: (1)..(75)
16 <220> FEATURE:
17 <223> OTHER INFORMATION: Coding region from residue 76 to 2406
18 <220> FEATURE:
19 <221> NAME/KEY: 3'UTR
20 <222> LOCATION: ()..(3977)
21 <220> FEATURE:
22 <221> NAME/KEY: polyA_signal
23 <222> LOCATION: ()..(4008)
24 <400> SEQUENCE: 1
25 gccgaggccc gcgcagtagc ggtactaagt agaggctgct ggacgcgccc caccgcgcac 60
26 caggcgggagc cagagatgct tgccagggcc gagcggcccc gcccgggccc ccggccgcct 120
27 ccggtctttc ctttcccgcc gccgctgtcg ctgctgctgc tgetggcgat actaagcgcc 180
28 ccggtgtgcg gccgcgtccc ccgctcagtg ccagaacct cgtgccccat ctccgaggct 240
29 gactcctatc tcacccggtt tgcagcgtct catacgtaca attactctgc tctccttggtg 300
30 gatcctgcct cccacacact ttacgtcggg gcacgggata gcatcttcgc tttaaccctc 360
31 cccttctctg gggaaagacc ccgaaggatc gactggatgg tacctgagac tcacagacag 420
32 aactgcagga agaaaggcaa gaaagaggac gaatgtcaca attttatcca gattctcgcc 480
33 attgtcaatg cctctcacct cctcacgtgc ggcaccttcg cttttgatcc gaagtgcggg 540
34 gttattgatg tgtccagttt ccagcaggtt gaaagacttg agagcggccg ggggaaatgt 600
35 ccttttgagc cagctcaacg gtcagcagct gtaatggctg ggggcgtcct ctacaccgcc 660
36 actgtgaaga acttctctgg gactgagccc atcatctccc gagctgtggg tcgagctgag 720
37 gactggattc gaacagagac cttgtcatcc tggcttaatg ctccagcctt tgtcgcagct 780
38 atggtcctga gccagctga gtggggggat gaagatggag acgatgaaat ctttttttcc 840
39 ttcacggaga cctcccgagt gttggactcc tatgagcgca tcaaggtccc aagagtggcc 900
40 cgagtgtgtg cgggggacct tgggggcagg aagacccttc agcagagatg gacgacgttt 960
41 ctgaaggctg acctgctgtg cccagggccg gagcatggcc gggcctccgg ggttctgcag 1020
42 gctatggcag agcttcggcc tcagcctgga gcgggaaccc ccattcttta tgggatcttt 1080
43 tcctcccagt gggaaggagc tgccatctct gctgtgtgtg cttccgacc ccaagacatc 1140
44 cgggcagtgcc tgaatggtcc ctttagagag ctaaaacatg actgcaacag gggactgcct 1200
```

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/284,180

DATE: 08/17/1999
TIME: 17:17:14

Input Set: I284180.RAW

```

45      gtcattggaca acgaggtgcc ccagcccaga cctggagagt gcatcgccaa caacatgaag 1260
46      ctccagcagt ttggatcctc actctccctg ccagaccgcg tgctcacctt tatcagagac 1320
47      caccctctca tggacaggcc cgtgttcccg gctgacggcc gcccctgct ggtcactaca 1380
48      gatacagcct atctcagagt cgtggccccc aggggtgacca gcctctcagg gaaagaatat 1440
49      gacgtgctct acctggggac agaggatgga cacctccacc gggctgtgcg cattggagct 1500
50      cagctcagtg tcttggagga tctggccttg tcccagaaac cacagccggt tgagagcatg 1560
51      aaattgtacc acgattggct cctgggtggc tcccatactg aggtgacaca agtgaacacc 1620
52      agcaactgtg gccgtctcca gagctgctcg gactgtatcc tggcccagga ccccggtgct 1680
53      gcctggagct tccggttga tgcctgtgtg gccacgccc gcgagcaccg cgggatgggt 1740
54      caagatatag agtcagcgga tgtctcttct ttgtgtccaa aagaacctgg agaacatccc 1800
55      gtagtgtttg aagttccggt ggctactgtg ggccacgtgg tccctgcatg tccccccagt 1860
56      tctgcctggg catcctgtgt gtggcaccag cccagtggag tgactgcgct cactccccgg 1920
57      agggatggac tagaggtggt ggtgacccca ggggcatagg gggcttatgc ttgcgagtgt 1980
58      caggagggtg gagccgccc cgtgggtggc gcttatagct tgggtgtggg cagccagcgg 2040
59      ggacctcaa accgggcccc caccgttgtg ggggctggat tgggtggctt tctcctgggt 2100
60      gttcttgcag catccctcac tctcctcctg attggtcgcc gtcagcagcg tcggcgacag 2160
61      agggagcttc tagctagaga caagggtggc ttagatctgg gggctccacc ttctgggacc 2220
62      acaagctata gtcaggacct tccctctcct tcgctgaag atgaacggct gccctggccc 2280
63      ctgggtaagc ggggcagtgg ttttgggtgg tccctccac ccttctgct ggattcttgc 2340
64      ccaagcccag cccacatccg gctcactggg gcgctctag ccacgtgtga tgagacctcc 2400
65      atctaaagcc ggggaaaatg actgccagcc atgagcagtc tctggaacta gtggctacca 2460
66      agaccatgat catggctgct cctttctctt ggagtctgtg tgttcacaca ttagtgtctg 2520
67      tctctggac ctggacctgg cctttgccc gatctctgat tctcatgaga gatcaacct 2580
68      gtaaccttct gcgatggcct cttgtcttgg gcccatcagc ttgtggggtg gagtaaggac 2640
69      ataggccccg gaaagggat cagtgtggag gtagttgggg cgtgtgtgcc ctgcgtcctt 2700
70      gtgggtggctg tatgatttcc cagtctgctg actctgggga gcgcatgac cctgactgc 2760
71      cttgagatct ctcccaactc agtttccctt tgcctggaa gactgtgtgt ctatacactg 2820
72      gtgtgcctag aaggcctgtc catgtgtgca tggacgacag ggccggtgcc tcgggtgctt 2880
73      tggggagtgc gagagaaagg ttggaatggg ggacaactta accctcggtg gccagtgagg 2940
74      gaaaccacat gcccgctccc atcacccac agcgcttctt taactttgag caaagtctcc 3000
75      aaagtgcact tctgggtggg aagggcagca ggacatgtgg ccccgctcct tctcctgtgc 3060
76      tttcccttct ggctgccaac cactgcctg ccaccgctgc gctttccctg gctggagtgg 3120
77      aggctgagtc ctctgtcctt ggtttccatt taaaatgaac ttcacaacat tctaaatatt 3180
78      gggggatgac aaatgacttt tttcccaga aaagtgtgta ggaaatacaa gcaggttaaa 3240
79      gaagatttgc ctcagtgact ttcacccttg ccctaaagca ggagtccctc agctagcgtc 3300
80      tgtggactcc ctgaaattgt atgcgtctgt ggactccctg aaattgtatg caaagtgtct 3360
81      gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtttgcgtgc atgtgtgcat 3420
82      gtgtgtttga tggctttcat cagattctca aggccttaat gaggttaaag gaccacggcc 3480
83      tatagtcacc acacttgggc cacatggagg aggtgttgc ctctgaggca gttcctcct 3540
84      ggctgctg aggcagccc ctggacacat tgctgctgga gacccacat ctctccagaa 3600
85      cttggaagct aggcctctgc cgtgcttgaa ggcaccacca tctcccttct tgcttcattc 3660
86      tctgtgtgct tctgctctg ctcatcctg ctcttggcct gtgaatgtgc ctgcccgtc 3720
87      cctgggtggg gacctcaaac cccagtgtgt atgctacct ttcagtgagg agtttctgtt 3780
88      ctgctttcct tgacagcagc ctgtgaacta ctacagagtc ccttgggttt ggagttcccg 3840
89      gtggctttga gtaggatctt tggcgtggca tctaacctag cagcattgat cgttcattgt 3900
90      aaagtgggga tatacctacc tcagggttgc tgcaaggatc aaacgaggaa acgtataaat 3960
91      aaagcattac ccacagcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 4008

```

92 <210> SEQ ID NO 2
93 <211> LENGTH: 2331
94 <212> TYPE: DNA

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/284,180

DATE: 08/17/1999
TIME: 17:17:14

Input Set: I284180.RAW

```

95 <213> ORGANISM: Rattus norvegicus
96 <220> FEATURE:
97 <223> OTHER INFORMATION: Coding region from residue 1 to 2331
98 <400> SEQUENCE: 2
99      atgcttgcca gggcgcgagcg gccccgcccc ggcccccggc cgccctccggt ctttcccttc 60
100     ccgcgcgcgc tgctgctgct gctgctgctg gcgatactaa gcgccccggt gtgcggccgc 120
101     gtcccccgct cagtgeccag aacctcgctg cccatctccg aggctgactc ctatctcacc 180
102     cggttttgcag cgtctcatac gtacaattac tctgctctcc ttgtggatcc tgccctccac 240
103     acacttttacg tcggtgcacg ggatagcatc ttcgctttaa ccctcccttc ctctggggaa 300
104     agaccccgaa ggatcgactg gatggtacct gagactcaca gacagaactg caggaagaaa 360
105     ggcaagaaaag aggacgaatg tcacaatttt atccagattc tcgccattgt caatgcctct 420
106     cacctcctca cgtgcggcac ctctgctttt gatccgaagt gcgggggttat tgatgtgtcc 480
107     agttttccagc aggttgaaaag acttgagagc ggccggggga aatgtccttt tgagccagct 540
108     caacggtcag cagctgtaat ggctgggggc gtccctctaca ccgccactgt gaagaacttc 600
109     ctggggactg agcccatcat ctcccgagct gtgggtcgag ctgaggactg gattcgaaca 660
110     gagaccttgt catcctggct taatgctcca gcctttgtcg cagctatggt cctgagccca 720
111     gctgagtggg gggatgaaga tggagacgat gaaatctttt ttttcttcac ggagacctcc 780
112     cgagtgttgg actcctatga gcgcacaaag gtcccaagag tggcccgagt gtgtgcgggg 840
113     gaccttgggg gcaggaagac ccttcagcag agatggacga cgtttctgaa ggctgacctg 900
114     ctgtgccagc ggcccgagca tggccggggc tccgggggtc tgcaggctat ggcagagctt 960
115     cggcctcagc ctggagcggg aacccccatc ttttatggga tcttttcttc ccagtgggaa 1020
116     ggagctgcca tctctgctgt gtgtgccttc cgaccccaag acatccgggc agtgctgaat 1080
117     ggtcccttta gagagctaaa acatgactgc aacaggggac tgccctgtcat ggacaacgag 1140
118     gtgccccagc ccagacctgg agagtgcacg gccacaaca tgaagctcca gcagtttgga 1200
119     tctcactctt cctgccaga ccgctgtctc acctttatca gagaccacct tctcatggac 1260
120     aggcctgtgt tcccggtga cgcccgcccc ctgctggtca ctacagatac agcctatctc 1320
121     agagtctgtg cccacaggtt gaccagctc tcagggaag aatatgacgt gctctacctg 1380
122     gggacagagg atggacacct ccaccgggct gtgcgcattg gagctcagct cagtgtcttg 1440
123     gaggatctgg ccttgttccc agaaccacag ccggttgaga gcatgaaatt gtaccacgat 1500
124     tggctcctgg tgggtcccca tactgaggtg acacaagtga acaccagcaa ctgtggccgt 1560
125     ctccagagct gctcggagtg tatectggcc caggaccccg tgtgcgcctg gagcttccgg 1620
126     cttgatgctt gtgtggccca cgccggcgag caccgcggga tggttcaaga tatagagtca 1680
127     gcggatgtct cttctttgtg tccaaaagaa cctggagaac atcccgtagt gtttgaagtt 1740
128     ccggtggcta ctgtgggcca cgtggtcctg ccatgttccc ccagttctgc ctgggcatcc 1800
129     tgtgtgtggc accagccagc tggagtact gcgctcactc cccggaggga tggactagag 1860
130     gtggtggtga cccaggggc catgggggct tatgcttgcg agtgtcagga ggggtggagcc 1920
131     gcccgctggg tggctgctta tagcttggtg tggggcagcc agcggggacc ctcaaaccgg 1980
132     gccacacccg ttgtgggggc tggattggtt ggctttctcc tgggtgttct tgcagcatcc 2040
133     ctactctcc tctgattgg tcgccgtcag cagcgtcggc gacagaggga gcttctagct 2100
134     agagacaagg tgggcttaga tctgggggct ccacctctg ggaccacaag ctatagtcag 2160
135     gacctccct ctccttcgcc tgaagatgaa cggtgcccc tggccctggg taagcggggc 2220
136     agtggttttg gtggcttccc tccaccttc ctgctggatt cttgcccaag cccagccac 2280
137     atccggtcga ctggggcgcc tctagccacg tgtgatgaga cctccatcta a 2331
138 <210> SEQ ID NO 3
139 <211> LENGTH: 776
140 <212> TYPE: PRT
141 <213> ORGANISM: Rattus norvegicus
142 <400> SEQUENCE: 3
143      Met Leu Ala Arg Ala Glu Arg Pro Arg Pro Gly Pro Arg Pro Pro Pro
144      1 5 10 15

```

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/284,180

DATE: 08/17/1999

TIME: 17:17:14

Input Set: I284180.RAW

145	Val	Phe	Pro	Phe	Pro	Pro	Pro	Leu	Ser	Leu	Leu	Leu	Leu	Leu	Ala	Ile
146				20					25						30	
147	Leu	Ser	Ala	Pro	Val	Cys	Gly	Arg	Val	Pro	Arg	Ser	Val	Pro	Arg	Thr
148			35					40					45			
149	Ser	Leu	Pro	Ile	Ser	Glu	Ala	Asp	Ser	Tyr	Leu	Thr	Arg	Phe	Ala	Ala
150		50					55					60				
151	Ser	His	Thr	Tyr	Asn	Tyr	Ser	Ala	Leu	Leu	Val	Asp	Pro	Ala	Ser	His
152		65				70					75					80
153	Thr	Leu	Tyr	Val	Gly	Ala	Arg	Asp	Ser	Ile	Phe	Ala	Leu	Thr	Leu	Pro
154					85					90					95	
155	Phe	Ser	Gly	Glu	Arg	Pro	Arg	Arg	Ile	Asp	Trp	Met	Val	Pro	Glu	Thr
156				100					105					110		
157	His	Arg	Gln	Asn	Cys	Arg	Lys	Lys	Gly	Lys	Lys	Glu	Asp	Glu	Cys	His
158			115					120					125			
159	Asn	Phe	Ile	Gln	Ile	Leu	Ala	Ile	Val	Asn	Ala	Ser	His	Leu	Leu	Thr
160		130					135					140				
161	Cys	Gly	Thr	Phe	Ala	Phe	Asp	Pro	Lys	Cys	Gly	Val	Ile	Asp	Val	Ser
162	145					150					155					160
163	Ser	Phe	Gln	Gln	Val	Glu	Arg	Leu	Glu	Ser	Gly	Arg	Gly	Lys	Cys	Pro
164					165					170					175	
165	Phe	Glu	Pro	Ala	Gln	Arg	Ser	Ala	Ala	Val	Met	Ala	Gly	Gly	Val	Leu
166				180					185					190		
167	Tyr	Thr	Ala	Thr	Val	Lys	Asn	Phe	Leu	Gly	Thr	Glu	Pro	Ile	Ile	Ser
168			195					200					205			
169	Arg	Ala	Val	Gly	Arg	Ala	Glu	Asp	Trp	Ile	Arg	Thr	Glu	Thr	Leu	Ser
170		210					215					220				
171	Ser	Trp	Leu	Asn	Ala	Pro	Ala	Phe	Val	Ala	Ala	Met	Val	Leu	Ser	Pro
172	225					230					235					240
173	Ala	Glu	Trp	Gly	Asp	Glu	Asp	Gly	Asp	Asp	Glu	Ile	Phe	Phe	Phe	Phe
174					245					250					255	
175	Thr	Glu	Thr	Ser	Arg	Val	Leu	Asp	Ser	Tyr	Glu	Arg	Ile	Lys	Val	Pro
176				260					265					270		
177	Arg	Val	Ala	Arg	Val	Cys	Ala	Gly	Asp	Leu	Gly	Gly	Arg	Lys	Thr	Leu
178			275					280					285			
179	Gln	Gln	Arg	Trp	Thr	Thr	Phe	Leu	Lys	Ala	Asp	Leu	Leu	Cys	Pro	Gly
180		290					295					300				
181	Pro	Glu	His	Gly	Arg	Ala	Ser	Gly	Val	Leu	Gln	Ala	Met	Ala	Glu	Leu
182	305					310					315					320
183	Arg	Pro	Gln	Pro	Gly	Ala	Gly	Thr	Pro	Ile	Phe	Tyr	Gly	Ile	Phe	Ser
184					325					330					335	
185	Ser	Gln	Trp	Glu	Gly	Ala	Ala	Ile	Ser	Ala	Val	Cys	Ala	Phe	Arg	Pro
186				340					345					350		
187	Gln	Asp	Ile	Arg	Ala	Val	Leu	Asn	Gly	Pro	Phe	Arg	Glu	Leu	Lys	His
188			355					360					365			
189	Asp	Cys	Asn	Arg	Gly	Leu	Pro	Val	Met	Asp	Asn	Glu	Val	Pro	Gln	Pro
190		370					375					380				
191	Arg	Pro	Gly	Glu	Cys	Ile	Ala	Asn	Asn	Met	Lys	Leu	Gln	Gln	Phe	Gly
192	385					390					395					400
193	Ser	Ser	Leu	Ser	Leu	Pro	Asp	Arg	Val	Leu	Thr	Phe	Ile	Arg	Asp	His
194					405					410					415	

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/284,180

DATE: 08/17/1999

TIME: 17:17:14

Input Set: I284180.RAW

```

195      Pro Leu Met Asp Arg Pro Val Phe Pro Ala Asp Gly Arg Pro Leu Leu
196                420                425                430
197      Val Thr Thr Asp Thr Ala Tyr Leu Arg Val Val Ala His Arg Val Thr
198                435                440                445
199      Ser Leu Ser Gly Lys Glu Tyr Asp Val Leu Tyr Leu Gly Thr Glu Asp
200                450                455                460
201      Gly His Leu His Arg Ala Val Arg Ile Gly Ala Gln Leu Ser Val Leu
202      465                470                475                480
203      Glu Asp Leu Ala Leu Phe Pro Glu Pro Gln Pro Val Glu Ser Met Lys
204                485                490                495
205      Leu Tyr His Asp Trp Leu Leu Val Gly Ser His Thr Glu Val Thr Gln
206                500                505                510
207      Val Asn Thr Ser Asn Cys Gly Arg Leu Gln Ser Cys Ser Glu Cys Ile
208                515                520                525
209      Leu Ala Gln Asp Pro Val Cys Ala Trp Ser Phe Arg Leu Asp Ala Cys
210                530                535                540
211      Val Ala His Ala Gly Glu His Arg Gly Met Val Gln Asp Ile Glu Ser
212      545                550                555                560
213      Ala Asp Val Ser Ser Leu Cys Pro Lys Glu Pro Gly Glu His Pro Val
214                565                570                575
215      Val Phe Glu Val Pro Val Ala Thr Val Gly His Val Val Leu Pro Cys
216                580                585                590
217      Ser Pro Ser Ser Ala Trp Ala Ser Cys Val Trp His Gln Pro Ser Gly
218                595                600                605
219      Val Thr Ala Leu Thr Pro Arg Arg Asp Gly Leu Glu Val Val Val Thr
220                610                615                620
221      Pro Gly Ala Met Gly Ala Tyr Ala Cys Glu Cys Gln Glu Gly Gly Ala
222      625                630                635                640
223      Ala Arg Val Val Ala Ala Tyr Ser Leu Val Trp Gly Ser Gln Arg Gly
224                645                650                655
225      Pro Ser Asn Arg Ala His Thr Val Val Gly Ala Gly Leu Val Gly Phe
226                660                665                670
227      Leu Leu Gly Val Leu Ala Ala Ser Leu Thr Leu Leu Leu Ile Gly Arg
228                675                680                685
229      Arg Gln Gln Arg Arg Arg Gln Arg Glu Leu Leu Ala Arg Asp Lys Val
230                690                695                700
231      Gly Leu Asp Leu Gly Ala Pro Pro Ser Gly Thr Thr Ser Tyr Ser Gln
232      705                710                715                720
233      Asp Pro Pro Ser Pro Ser Pro Glu Asp Glu Arg Leu Pro Leu Ala Leu
234                725                730                735
235      Gly Lys Arg Gly Ser Gly Phe Gly Gly Phe Pro Pro Pro Phe Leu Leu
236                740                745                750
237      Asp Ser Cys Pro Ser Pro Ala His Ile Arg Leu Thr Gly Ala Pro Leu
238                755                760                765
239      Ala Thr Cys Asp Glu Thr Ser Ile
240                770                775

```

241 <210> SEQ ID NO 4

242 <211> LENGTH: 2315

243 <212> TYPE: DNA

244 <213> ORGANISM: Homo sapiens

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

PAGE: 6

VERIFICATION SUMMARY
PATENT APPLICATION US/09/284,180

DATE: 08/17/1999
TIME: 17:17:14

Input Set: I284180.RAW

Line ? Error/Warning

Original Text

481 W "N" or "Xaa" used: Feature required

Xaa Asp Tyr Cys Xaa Trp